Preferences over equality in the presence of costly income sorting

Gilat Levy and Ronny Razin, LSE

Abstract: We analyze preferences over redistribution in societies in which there are complementarities in income and agents use costly signals to sort themselves according to income. We characterize conditions over income distributions which imply that the median voter will prefer full redistribution to an environment in which he is able to match, at a cost, with agents with higher income. We relate these conditions to income inequality as well as to the properties of increasing or decreasing failure rates, which are commonly used to approximate real income distributions. When we consider only local changes, we illustrate how an "ends against the middle" coalition of voters might arise to increase the exclusiveness of sorting.

1   Introduction

Expensive private schools, large houses, charity donations, extravagant weddings and designer clothes, among others, have all been considered in the literature as ways in which individuals try to signal their wealth to others.[1] In environments in which there are complementarities in wealth, individuals have an incentive to use such tools in order to identify similar ones. Thus, beyond being a traditional tool for creating equality, policy instruments such as income or wealth redistribution have an additional effect in such societies; they reduce the incentive to signal as well as the occurrences of sorting.

   While poor income groups are naturally against income signalling which allows others to avoid matching with them, middle income groups may be particularly affected by the pressure to signal. As inequality increases, such groups might be under increased pressure to interact with wealthier groups which may put a large strain on their budgets. In this paper we explore how this incentive to signal shapes individual preferences over redistributive policies and whether agents such as the median voter prefers a fully equal society or a society in which he can mix with the rich, but at a cost.

   We analyze a simple model of linear redistributive taxation in which individuals have a utility that is supermodular in their own income and the income of the individuals they interact with.

---

[1] The literature on conspicuous consumption includes contributions by Liebenstien (1950), Bagwell and Bernheim (1996), Pesendorfer (1995) and Heffetz (2011). Glazer and Konrad (1996) consider signalling of wealth via charitable donation which exhibits positive externalities. Moav and Neeman (2010) analyze the trade-off between conspicuous consumption and human capital as signals for unobserved income and show why the poor spend a large share of their income on conspicuous consumption.

Individuals match randomly with (and only with) those who acquire the same costly (and wasteful) signal, and we focus on a partition of the income distribution into such finite number of groups or clubs which are determined by a set of costly signals and the respective incentive compatability constraints.

We characterize the relation between the income distribution function and the political viability of redistributive policies. We start by comparing societies with signalling/sorting to fully equal societies. For distributions for which the median income equals the mean, in the absence of sorting, the median voter has no strong preferences for taxation. We are then able to find a necessary and sufficient condition for the median and all those below to prefer full redistribution to *any* form of coarse signalling. This condition is more likely to hold when the income distribution is relatively equal. In other words, when the median has no incentives for redistribution in the absence of sorting, such incentives arise in the presence of sorting when the underlying income distribution is sufficiently equal.

For skewed distributions for which the median income is less than the mean income, we focus on properties that relate to the hazard rate which are commonly used to approximate real income distributions.[2] When the distribution function exhibits an increasing failure rate (IFR), any form of signalling is neither efficient nor politically viable compared with full equality.[3] In particular, in this case all those with incomes below the mean income (and some with income above) prefer full equality to any combination of taxation and signaling. Moreover, we show that for an open set of distributions with decreasing failure rate (DFR), any form of signalling, although efficient, is not politically viable. The result illustrates that once the income distribution is skewed and the median has redistribution motives even in the absence of sorting, such incentives also arise in the presence of sorting when the income distribution is sufficiently equal and when it is sufficiently unequal.

We next analyze what happens when only small changes can be made to redistributive policies. We show that it is sometimes the case that the median voter and all those above prefer to reduce taxation. This arises if sorting is sufficiently inclusive, and if, again, society is sufficiently unequal. Finally, we consider policies that can affect the exclusiveness of sorting, e.g., subsidies or taxes for private schools. We show how an "ends against the middle" coalition might arise in which the middle class faces opposition from both sides to increasing the inclusiveness of sorting and decreasing its price.

---

[2] See for example Singh and Maddala (1976) and Salem and Mount (1974).

[3] We in fact characterize a more general condition (NBUE) which is always satisfied by an IFR distribution. Note that as the utility from signalling is convex, whenever sorting is inefficient compared with full equality, the mean and thus all those below will also prefer full equality to sorting.

Our paper is in general related to the political economy literature on preferences over redistribution, such as Meltzer and Richards (1981). While they highlight a channel that may decrease the preferences for redistribution -e.g., tax distortions or reduction in labour incentives- we highlight a channel that can increase the preferences for redistribution of agents above the median and sometimes even above the mean. This is consistent with the observation that along with poor individuals who vote to parties on the right, which had received much attention in the literature,[4] there are also rich individuals voting left. For example, De la O and Rodden (2008) use the Eurobarometers and World Values Survey data to show that on average well over 40% of the wealthiest individuals vote for parties of the left in Europe.

Our model is closely related to recent literature on the cost of signalling. Hoppe, Moldovanu and Sela (2009) consider a model in which individuals signal their attributes. Their model is an incomplete information model with two sided heterogeneity, finite types, and perfect signalling. Among other questions, they ask whether costly signalling provides higher average welfare compared with random matching.[5] We discuss the relation of our results to theirs when we analyze the benchmark of efficiency. Several other papers focus on coarse matching, for example Hoppe, Moldovanu and Ozdenoren (forthcoming) and McAfee (2002), and show the conditions under which coarse matching provides sufficiently high surplus compared with random or perfect matching. While all these papers consider efficiency, we consider individual welfare and specifically political viability. Following Hoppe, Moldovanu and Sela (2009) and Hoppe, Moldovanu and Ozdenoren (forthcoming) we also use results from reliability theory from Barlow and Proschan (1966).

Finally, our paper is related to the literature on sorting in the tradition of Tiebut models, where agents who have different preferences over the provision of public goods sort themselves into communities which decide via majority rule on the level of such provision.[6] Within this literature several papers consider the effect of redistributive policies. Fernandez and Rogerson (2003) consider provision of quality of schooling and analyze different equalizing policies which target the finance of education. Epple and Romano (1998) model the supply side, i.e., the market for private schools, and show how more wealthy and able agents are screened into better quality schools. In this environment, they consider the policy of school vouchers and show that it is mainly high ability and high income types who benefit from the introduction of vouchers to private schools. Fernandez and Gali (1997) analyze whether markets or tournaments produce a more efficient outcome, and show that with credit constraints, markets perform less well than

---

[4]See for example Roemer (1998).

[5]See also Rege (2003).

[6]See Fernandez and Rogerson (2001).

3

tournaments at sorting individuals according to ability.[7]

The remainder of the paper is organized as follows. We present the model in the next Section. In Section 3 we analyze conditions for the median to prefer full equality to all forms of finite sorting for distributions for which the median equals the mean income. In Section 4 we look at more general distributions and establish conditions both for efficiency and for political viability. Section 5 considers local preferences over more smooth taxation policies and over changing the exclusiveness of the club. An appendix contains all proofs.

2   The model

Suppose that agents differ in their income, x, which is distributed according to some $F(x)$ and density $f(x)$ (positive everywhere) on some $[0, \nu]$, $\nu \leq \infty$.[8] Let μ (x^m) denote the mean (median) of the distribution. We will consider income distributions with $x^m \leq \mu$. Agents are matched with one another (a process which we will describe shortly); when two agents x and y meet, each enjoys $u(x, y) = xy$. We therefore assume complementarities in income, as in marriage or partnership markets or as in peer benefits from education (for which income can potentially induce a higher quality).[9] Note that for simplicity we focus on one dimension of heterogeneity. Our analysis can be generalized to consider two dimensional heterogeneity, of for example ability and income.[10]

We consider a matching process in which agents might wish to signal their income to one another. Naturally, given the complementarities above, agents with high income would like to differentiate themselves from the rest and match with one another and might therefore use costly signals to do so. When no signalling devices arise, matching is random. When some agents use a costly signal, they will match randomly with, and only with, all agents who use the same signal. Agents who do not acquire any signal are randomly matched with all others who do not acquire any signal. When an agent uses some signal that costs b, his utility will

---

[7]In a more complex environment, Benabou (1996) analyzes a dynamic model and looks at the effect of stratification on growth and efficiency.

[8]The lower bound of the support is a normalization. Indeed, one example we discuss below is the Pareto distribution on $[1, \infty)$.

[9]The analysis can be extended to other more generalized utility functions of the form $h(x)g(y)$ exhibiting supermodularity with the conditions on F adjusted. We use a simple form of complementarity for tractability and in order to focus on the properties of F. See also Remark 1.

[10]Note that in papers which take into consideration multidimensional heterogeneity, where for example agents differ in both ability and income, such as Fernandez and Rogerson (2003) or Epple and Romano (1998), it is typically the case that a single crossing property is assumed which implies that agents sort themselves according to income.

4

be $x_i E[x_j | j \in X_b] - b$ where X_b is the set of other agents who also use the costly signal b. The quasi-linear nature of the utility function is simple to use but is not necessary for our results; our results hold if for example the utility of an agent x_i who matches with an agent x_j is $(x_i - b)(x_j - b)$ instead (see Remark 1).

By single crossing, if some agent with x_i prefers a signal b over b', all agents with $x > x_i$ will prefer b over b'.[11] We will therefore focus on monotone sorting, i.e., with connected intervals. We will abstract away from the supply side, i.e., how signals such as private schools or real estate prices are determined.[12] But as agents are assumed to choose optimally whether to acquire a costly signal, no matter how the supply side arises, the costs of the signals must satisfy some incentive compatability constraints. Suppose that there is only one such signal, where all agents above some cutoff x acquire this signal and pay $b(x)$ and all below pay nothing. The type at the cutoff x will be indifferent and hence the price of the signal must satisfy:

$$b(x) = x(\bar{E}_x - \underline{E}_x)$$

where

$$
\begin{aligned}
\underline{E}_x &= E[v | v \le x] = \frac{\int_0^x v f(v) dv}{F(x)} = x - \frac{\int_0^x F(v) dv}{F(x)} \\
\bar{E}_x &= E[v | v \ge x] = \frac{\int_x^\nu v f(v) dv}{1 - F(x)} = x + \frac{\int_x^\nu (1 - F(v)) dv}{1 - F(x)}
\end{aligned}
$$

The expected utility of an individual $x' < x$ is therefore $x' \underline{E}_x$ and the expected utility of an individual $x' > x$ can be written as:

$$
\begin{aligned}
& x' \bar{E}_x - b(x) \\
=\ & x' \bar{E}_x - x(\bar{E}_x - \underline{E}_x) \\
=\ & (x' - x)\bar{E}_x + x\underline{E}_x
\end{aligned}
$$

Expected utility can be interpreted as the utility of the cutoff type, plus an information rent component that depends on the distance from the cutoff. The utility from signalling is, as usual, increasing and convex in x'. More generally, below we define the feasible signal partitions that we consider:

---

[11] If prices do not differ, we can sustain non monotone signalling structure for example when top and bottom agents use the same signal and middle agents use another signal, with the same expectations over income in both cases. For each individual, this is equivalent in expectations to all society being matched randomly and we will treat it as one signal.

[12] For such analysis see Damiano and Li (2007) and Rayo (2005).

Definition 1: *A feasible signal partition (FSP) is a vector* $\mathbf{x} = (x_0, x_1, ... x_{n-1}, x_n)$ *with* $x_0 = 0$ *and* $x_n = 1$ *and* $x_i \leq x_{i+1}$, *such that all agents with type* $x \in [x_i, x_{i+1}]$ *for* $i = 0, 1, ..., n-1$ *pay* b_i *and are matched randomly with agents in* $[x_i, x_{i+1}]$ *only, with*

$$b_0 = 0$$
$$b_i - b_{i-1} = x_i(E[x_j | x_j \in [x_i, x_{i+1}]] - E[x_j | x_j \in [x_{i-1}, x_i]]).$$

Note that in the formulation above we are restricting the price of joining the lowest element in the partition to be zero. As our results are mainly about showing when sorting is not politically viable this assumption strengthens our results.

In that environment, i.e., given some FSP, we ask what are the preferences of the median voter for redistribution (the preferences of the median in our environment will be sufficient to represent a majority). We will look at a simple linear taxation scheme in which the disposable income of an agent of type x is $x(1-t) + t\mu$. The key premise that is built into the analysis is that when income inequality is reduced, so are the incentives to sort or the willingness to pay for sorting. As long as the absolute after-tax level of income has some effect on the quality of the match or on the incentives to sort, our results will qualitatively hold.

For most of the paper we will consider the median voter's preferences for full redistribution (FR); the utility from full redistribution is μ^2 and is equal to all (In Section 5 we consider local preferences over redistribution). Moreover, we will look at conditions on F for which the median will prefer FR *irrespective* of what the FSP is. That is, a condition that allows us to know that FR garners enough electoral support without any information of how society is divided into special clubs or what the tax is.

Note that in the absence of sorting, preferences over redistribution in the model are "standard": all agents up to the mean will prefer redistribution, and all those below will be against redistribution. In the presence of perfect sorting, the median is against sorting. In that case, all are matched with their own type but have to pay a price for it; thus, even the mean in the population would rather equalize income in society, in which case he also (trivially) matches with his own type but need not pay a price.[13] When sorting is coarse or discrete though, which is likely to be the case in reality, preferences are not as clear cut; discrete signalling will leave more rent to the agents. Thus, if there is any signalling structure that will be politically viable, it is more likely to be coarse which is why we focus on such structures.

---

[13]Specifically, the cost of signaling for type x will be $b(x) = x^2/2$ and so the type at that is indifferent is $x' = \sqrt{2}\mu > \mu$.

3 A simple condition when no redistribution motive exists

Within coarse signalling, it is naturally the case that the poor who are left behind and do not belong to a "club", prefer FR as they can then match with better income types. Similarly, the very rich might prefer sorting as their utility from this can outweigh the cost which is determined according to the IC constraint of a lower income type. But the middle classes' preferences are not obvious: on the one hand they can match with higher income types, whereas on the other hand the price for this compared with their willingness to pay is relatively high.

We first look at income distributions for which the mean income is equal to the median income. For such distributions, in the absence of signaling, the median has no strong preferences for or against redistribution. Therefore, any such preference will be driven by the effects of signalling.

Consider first a simple FSP with $n = 2$ (i.e., just one cutoff x) and $t = 0$. Clearly, FR is preferred by the median if the cutoff x is such that $x > x^m$, i.e., when he does not belong to the club. FR in this case will increase both his utility and the utilities of those he interacts with. We therefore have to focus on signals that satisfy $x \leq x^m$. The median prefers full redistribution to any signal x iff

$$(x^m - x)\bar{E}_x + x\underline{E}_x \leq \mu^2$$

Plug $x^m = \mu$ and divide by μ to get

$$(1 - \frac{x}{\mu})\bar{E}_x + \frac{x}{\mu}\underline{E}_x \leq \mu$$

As

$$\mu = (1 - F(x))\bar{E}_x + F(x)\underline{E}_x, \tag{1}$$

then FR is preferred to coarse sorting for any x iff:

$$\frac{x}{\mu} \geq F(x) \text{ for all } x \leq \mu \tag{Condition 1}$$

It is easy to see that Condition 1 relates to income inequality: For any x, fixing μ, if there is a higher weight on incomes below x then the condition is less likely to hold. Thus, any mean preserving spread of a distribution that does not satisfy the condition will not satisfy it as well, and any mean preserving contraction of a distribution which satisfies the condition will do so as well. As another illustration consider all distributions which are symmetric around the mean; all with density functions which are inverted U-shaped satisfy Condition 1 and all that have U-shaped functions satisfy it as long as they are not too concave (a necessary and

sufficient condition being that $f(0) < \frac{1}{\mu}$). Also, any shift in the distribution in a first order stochastic sense implies that the condition is more likely to hold. Thus, potentially, dynamics can play a role: Sorting is more likely to occur when the distribution is already unequal, which might increase inequality further in the long run.

Example 1: Consider the family of the symmetric beta distributions $f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}du}$ on $[0, 1]$ with $\alpha = \beta$. For this family the Gini coefficient is monotonically decreasing in α. When $\alpha > 1$ the density is inverted U-shaped, and Condition 1 holds. When $\alpha < 1$ Condition 1 is not satisfied for an interval of $x's$; the distribution function is then bimodal with the modes close to zero and one. A low cutoff x guarantees that the median (at 0.5) stays away from a relatively large mass of low types, as well as a low cost of signalling, and thus the median is in favour of sorting compared with FR.

Generalizing the above to any level of taxation t is straightforward. Generalizing Condition 1 to any FSP with $n > 2$ does not however follow immediately. In particular, whenever F is concave, adding more signals below some cutoff x reduces the signalling cost for all types $x' > x$ and thus improves the utility from signalling (see Lemma 1). Still, we are able to show that Condition 1 is necessary and sufficient for all FSP, where in the proof we use induction and Condition 1 repetitively.

Proposition 1: *The median prefers FR to any FSP and any tax level t iff F satisfies Condition 1.*

Remark 1: When agents match in our model their utility from the matching incomes excludes the payment of the signal. Alternatively, one can assume that the utility from the match is $(x_i - b)(x_j - b)$. We show in the appendix that Condition 1 is sufficient in this case for the median to prefer FR. For other more general utility functions, for example if an agent enjoys his income on top of the utility from matching, such as in $u(x_i) - b + h(x_i)g(x_j)$, it is possible to construct an adjusted condition.

4 Skewed distribution functions

We now consider more general distribution functions for which $x^m \leq \mu$. It will be useful to start with efficiency analysis (average utility) to compare between FR and any FSP, to pinpoint some relevant properties of income distribution functions.

4.1 Efficiency

Again, to gain intuition, we start with one signal x (and $t = 0$). Aggregate utility from full redistribution is $U(FR) = \mu^2$, and aggregate utility from sorting is

$$
\begin{aligned}
U(x) &= \underline{E}_x \int_0^x v f(v) dv + (1 - F(x)) x \underline{E}_x + \bar{E}_x \int_x^\nu (v - x) f(v) dv \\
&= F(x) \underline{E}_x^2 + (1 - F(x)) \bar{E}_x^2 - x(1 - F(x))(\bar{E}_x - \underline{E}_x)
\end{aligned}
$$

Using (1) and re-arranging, we have that

$$
\begin{aligned}
&U(x) - U(FR) \\
&= F(x) \underline{E}_x(\underline{E}_x - \mu) + (1 - F(x)) \bar{E}_x(\bar{E}_x - \mu) - x(1 - F(x))(\bar{E}_x - \underline{E}_x)
\end{aligned}
$$

The first two elements represent the potential benefit from sorting vs. FR, and the last element is the cost of signalling. Using (1) again to plug for $(\underline{E}_x - \mu)$ and $(\bar{E}_x - \mu)$, we have that

$$
U(x) - U(FR) < 0 \Leftrightarrow \bar{E}_x - \underline{E}_x < \frac{x}{F(x)}
$$

Intuitively, efficiency is related to the spread of the expectations at the top group vs. the bottom group given x, compared to the cost of signalling which is proportional to x. If the spread is large enough compared to $\frac{x}{F(x)}$, then sorting is efficient. As $F(x)(\bar{E}_x - \underline{E}_x) = \bar{E}_x - \mu$, we get that the above condition is equivalent to

$$
\bar{E}_x - \mu < x
$$

Note that $\bar{E}_x - x = \frac{\int_x^\nu (1 - F(v)) dv}{1 - F(x)}$. In Renewal Theory, the property that $\frac{\int_x^\nu (1 - F(v)) dv}{1 - F(x)} < \int_0^\nu (1 - F(v)) dv = \mu$ is called NBUE (new better than used in expectations), whereas the property $\bar{E}_x - x > \mu$ is called NWUE (new worse than used in expectations). Hall and Wellner (1984) showed that any NBUE function has a coefficient of variation $CV(x) = \frac{\sqrt{Var(x)}}{E(x)} \leq 1$, whereas for any NWUE, $CV(x) \geq 1$. For the case of perfect continuous signalling, Hoppe, Moldovanu and Sela (2009) show that $CV(x) \geq (\leq)1$ is a sufficient and necessary condition for sorting to be efficient (not efficient). Intuitively, in the limit, the cost of signalling is proportional to the level of output while the benefit of signalling increases in heterogeneity. For the discrete model though, the cost of signalling is still proportional to output differences and thus the condition on the coefficient of variation is necessary but not sufficient.[14] Although the proof is more complicated, we can generalize the above condition for any FSP:

---

[14]Hoppe, Moldovanu, and Sela (2009) look at the utility from random matching but at the aggregate this equals the utility from FR to the whole of society. For their discrete model which has incomplete information on a discrete set of types but perfect signalling, a necessary and sufficient condition for efficiency (inefficiency) of signalling is for the function to be DFR (IFR).

Proposition 2: *Compared with FR, sorting is efficient (not efficient) for any FSP iff F is NWUE (NBUE).*

Barlow and Proschan (1966) have shown that any function with a decreasing failure rate (DFR) -such as the exponential, Pareto, Weibull and Gamma (for shape parameter less than one)- is also NWUE, and any function with an increasing failure rate (IFR) -such as exponential, uniform, normal, Weibull and Gamma (for shape parameter greater than one)- is also NBUE. Thus a sufficient condition for the efficiency of sorting can be presented in terms of failure rates.[15]

Corollary 1: *Sorting is efficient (not efficient) for any FSP if F has decreasing (increasing) failure rate.*

Focusing on IFR/DFR distribution is useful as some of them are commonly used and we will use these properties for the analysis of political viability below; there is in fact a substantial literature on fitting income distributions to real data which relates to the failure rate properties, which measure the odds against advancing to a higher income given a particular income level is reached. Salem and Mount (1974) have advocated a version of the Gamma distribution, which is IFR, where $f(x) = \frac{\lambda^{\alpha}}{A(\alpha)} x^{\alpha-1} e^{-\lambda x}$ on $[0, \infty]$ for $A(\alpha) = \int_0^{\infty} e^{-u} u^{\alpha-1} du$.[16] Other distributions which are considered in the literature are Pareto (which is DFR) and the Lognormal (which is first IFR and then DFR). Singh and Maddala (1976) claim that income distributions should be DFR at least for high enough income, as the ability to make more money should increase with one's income, once some threshold is reached.[17]

4.2 Political viability

We now consider the preferences of the median voter over signalling. Given (1), we have that

$$\bar{E}_x - \mu \leq x \Longrightarrow \mu < \frac{x}{F(x)}$$

and hence NBUE implies Condition 1. This is obvious as whenever sorting is inefficient, due to the convexity of utility from sorting, also the mean and *a fortiori* the median will prefer FR.

---

[15] See also Hoppe, Moldovanu and Ozdenoren (forthcoming) and Hoppe, Moldovanu and Sela (2009).

[16] For this distribution the median is $\frac{3\alpha-1}{3\lambda}$, $\frac{1}{\sqrt{\alpha}}$ is the parameter of skenewss, and the mean is $\frac{\alpha}{\lambda}$. For the decades of the 60's, their estmiate of α is around 2 and λ is around 0.03.

[17] Singh and Maddala (196) fit the data to some mixture of Pareto and Weibull, with an increasing *proportional* hazard rate $(x \frac{f(x)}{1-F(x)})$ which then converges to become constant. We note that Cramer (1978) advocates caution with respect to interpreting failure rates properties with regard to income (where such properties should relate to time or age).

Corollary 2: *Whenever F is NBUE, for any FSP and tax t, sorting is not politically viable compared with FR. In particular, the coalition in favour of FR includes all those with incomes below or equal to the mean income.*

Given the above, it will be instructive to focus on F for which sorting is efficient compared with FR and concretely we focus on functions which exhibit DFR.

First note that a tighter version of Condition 1 (whose proof is analogous, see the proof of Proposition 1 in the appendix) will demand that $\frac{x}{F(x)} > x^m$ for all $x < x^m$. This condition implies that any FSP (and any t) is dominated for the median by random matching (which provides utility $x^m \mu$), which in turn is dominated by FR.

Note though that any DFR function is also concave and if F is sufficiently concave, it will surly violate $\frac{x}{F(x)} > x^m$ for small enough x. Also, as this condition does not take into account the benefits of redistribution, we need to specify a tighter condition. The following Lemma will be helpful:

Lemma 1: *(i) Suppose that $z \in [x_i, x_{i+1}]$ where $[x_i, x_{i+1}]$ is an element of an FSP vector* **x***, and consider any FSP* **x**$'$ *that identifies with x on $[0, x_{i+1})$ but is more coarse on $[x_{i+1}, 1]$. The utility of z is higher for any such* **x**$'$ *and is maximized when $x_{i+1} = 1$. (ii) Suppose that $z \in [x_i, x_{i+1}]$ where $[x_i, x_{i+1}]$ is an element of an FSP vector* **x***, and consider any FSP* **x**$'$ *that identifies with x on $[x_i, 1]$ but is finer on the region $[0, x_i)$. When F is concave, the utility of z is larger from* **x**$'$ *and converges in the limit (when the partition below x_i becomes perfect) to*

$$\frac{x_i^2}{2} + (z - x_i) E[x_j | x_j \in [x_i, x_{i+1}]].$$

Whenever F is DFR and thus concave, to see that any FSP is inferior to FR from the point of view of the median, we can therefore focus on the limit of continuous perfect signals below some $x < x^m$, and no other signals above x. For any x, such a signalling structure provides the maximum utility for the median, which is (when $t = 0$):

$$\frac{x^2}{2} + (x^m - x)\bar{E}_x \tag{2}$$

Note that for the utility in (2) to be higher than FR, x has to be high enough; otherwise, for too small x, $\bar{E}_x \to \mu$ and the best the median can get from sorting approaches his utility from random matching. But also to provide information rent, x cannot be too high, which implies that political viability of sorting is tricky. We then find:

Proposition 3: *Suppose that F is DFR and let x^* satisfy $\bar{E}_{x^*} = \frac{\mu^2}{x^m}$. The median prefers FR to any FSP and any t if F satisfies one of the following conditions: (i) $x^* > x^m$; (ii) $\frac{x^m}{\mu} \le \frac{1}{2}$; (iii) $\frac{f(x^*)}{1-F(x^*)} < \frac{1}{x^m-x^*}$.*

Example 2: The Pareto distribution on $[1,\infty]$ is DFR and satisfies (i), i.e., for all shape parameters we have that $x^* > x^m$.

Note that for all DFR's, $\frac{x^m}{\mu} \le \ln 2 \approx 0.69.$[18] Thus condition (ii) above covers a large set of DFR's and in particular those that are relatively more concave or more unequal. In this case the information rent is simply too small for the median to gain enough utility in any sorting environment. Condition (iii) arises when f falls fast enough so that the hazard rate is low enough at x^* which again implies a high degree of concavity or income inequality. Intuitively, this implies that the utility from sorting decreases with x as only little is gained in terms of the expectations over the high income types above x when x increases, as f is already sufficiently flat. On the other hand if F is not sufficiently concave, and does not increase fast enough to rid of large weight on small values, $\bar{E}_x$ will not increase quickly enough implying that it might be that $x^* > x^m$.

Note that the analysis conducted in Section 3 focused on an environment in which the median has no pure redistribution motive, in which case inequality implied that it is more likely that the median prefers sorting. Once redistribution motives are introduced, both a sufficient degree of inequality and a too little degree of inequality imply that the median is against sorting. The conditions together impose therefore quite tight restrictions and in fact we have not been able to find any example of a DFR function for which there is some sorting which is politically viable.

Remark 2: What about distributions which are neither IFR nor DFR? Consider the log-normal distribution, which is first IFR and then DFR. When for example the median equals 1, if σ is not too high (i.e., if F is not too concave), then $\frac{x}{x^m} > F(x)$ which implies that the median prefers sorting to FR. Another example is $F(x) = x^\alpha$ on $[0,1]$ which, for $\alpha < 1$, is first DFR and then IFR; if $\alpha < 0.35$ then $x^* > x^m$. On the other hand recall the symmetric beta distribution (Example 1) with $\alpha = \beta = 0.5$. This distribution is neither IFR or DFR. The density is bimodal with most weight on the tales, and the mean and the median equal 1/2. For this distribution, we have that $\frac{1}{2} > \frac{x}{F(x)}$ for x small enough which implies that the mean

---

[18]By Theorem 4.7 in Barlow and Proschan (1965), if F is DFR then $F(x) \ge 1 - e^{-\frac{x}{\mu}}$ for any $x < \mu$. This implies that all DFR with the same μ as some exponential, have a lower median, and in particular that $\frac{x^m}{\mu} \le \ln 2$ which is the ratio between the median and the mean in the exponential.

and the median prefer signalling with such a cutoff x to FR (and hence it is also efficient). An interesting point to note is that if a monopoly would choose x to maximize profits (e.g., $b(x)(1 - F(x))$), it would choose $x = 0.5$ and thus implementing this politically viable and efficient FSP is not trivial.

5 Preferences for small reforms in redistribution and sorting.

As we have seen, for many income distribution functions, the median prefers to have a completely equal society rather than live in a society in which he has to pay to interact with high income types. We now consider two extensions. First, we consider local preferences over linear taxation and show that the median (and all those above) might prefer taxes to be reduced and income inequality to grow. This happens when the income distribution is sufficiently unequal. Second, there are other policy tools that might affect the equilibrium in the signaling market; the government can introduce constraints or subsidies in the housing or education markets for example, which will affect the price and composition of sorting. We will show that in this case an "ends against the middle" coalition can arise between high income types in the club and low income ones outside the club, to increase the exclusiveness of the club. We focus in this section on a simple FSP with $n = 2$ and one cutoff x.

With regard to more smooth taxation policies, consider linear taxation where the income of an individual i is $(1 - t)x_i + t\mu$. For some type $x' > x$, the utility from some cutoff x can therefore be written as

$$((1 - t)x_i + t\mu)(\underline{E}_x(1 - t) + t\mu)$$
$$+(x' - x)(1 - t)(\bar{E}_x(1 - t) + t\mu)$$

Taxation increases the utility of all agents who are not in the club, as long as $x < \mu$. For agents in the club, it increases the base utility of the agent at the cutoff, but decreases the gain from the information rent, and the income of those whom they are matched with. This implies that if the cutoff type is low enough, and taxation is low enough as well, the first positive effect is very small. Thus the median actually prefers to reduce it further as the second negative effect dominates the first one. We can then show:

Proposition 4: *When local changes only are possible to taxation, then when the club is relatively inclusive and taxation is low enough, the median and all those above support reducing taxation. The median and all those below support an increase in taxation if it is sufficiently high already.*

An implication of Proposition 4 is that the incentive for taxation can be lower under sorting than in its absence, if taxation is low enough and the club large enough, as in the absence of sorting the median always favours taxation. When taxation is already high enough, or society relatively equal, the negative effect of taxation on the information rent is rather small and is outweighed by the positive effect on the utility of the type at the cutoff. In fact, the incentive to tax is then greater in the presence of sorting than in its absence. Once again, we see that inequality may lead to reduced taxation and hence more inequality whereas equality may lead to increased taxation and hence to more equality.

Next we analyze local preferences over x; will agents prefer the club to be more or less inclusive? For the poor who are not in the club, the higher is x the better is the average income for their match. For those in the club, the derivative of the utility from sorting is (for some type x') is:

$$(\bar{E}_x - x)((x' - x)\frac{f(x)}{1 - F(x)} - 1) + (x - \underline{E}_x)(x\frac{f(x)}{F(x)} - 1) \tag{3}$$

An increase in x increases $\bar{E}_x$, $\underline{E}_x$, as well as the price. What is clear from (3) however, is that once x' prefers an increase in x, then all those above prefer an increase in x as well. This reveals a possible "ends against the middle" coalitions for small local changes.

Proposition 5: *A coalition to increase x will always consist of agents below x and sometimes consists of all agents from some $x' > x$ and above. Moreover, there exist income distributions for which an "ends against the middle" coalition can arise to successfully increase the exclusiveness of the club.*

Policies that can increase or decrease x are subsidies or taxes imposed on the signalling devices; a fuller analysis will naturally include the cost of taxation. Still, the result above indicates who may gain and who may lose from such a policy. We prove that such successful "ends against the middle" coalitions (consisting of more than 50% of the population but excluding the median) can arise by examples:

Example 3: Consider the Gamma distribution as in Salem and Mount (1974) with the parameters they estimate for the income distribution in the US in the 1960's, $\alpha = 2$ and $\lambda = 0.03$. For these parameters, $x^m \approx 55$. When $x = 40$, all types with income above 96 prefer to increase x, together with all types below 40, composing a share greater than half the population.

Example 4: Consider the exponential distribution with $\lambda = 2$ where $x^m = 0.346$ and the

mean is 0.5. For $x = 0.25$, then all types above 0.78 and all types below 0.25 would rather increase x which comprises a coalition of 60%.

Example 5: Consider the uniform distribution over $[0, 1]$ and assume some linear tax t and a cutoff x. For all $x' \geq 1 - 0.5\frac{t}{1-t}x$, an increase in x increases utility. Thus an "ends against the middle" coalition for an increase in x is greater than 50% of the population whenever $x \geq \frac{0.5}{1+0.5\frac{t}{1-t}x}$. For example, when $t = 0.5$, this is the case whenever $x \geq \frac{1}{3}$.

6 Discussion

Our analysis implies that when preferences for matching and signaling are taken into account, they can affect preferences regarding taxation, and that in most cases, such preferences become stronger for the poorer majority of the population. In fact, we couldn't find an example with a DFR distribution in which the median voter preferred sorting to full equality. We have also shown how in some cases agents with income at the mean will strictly favour redistribution (when Condition 1 is strictly satisfied as is the case with an IFR distribution). This is consistent with the observation that along with poor voters voting to the right, which had received much attention in the literature, there are also rich voters voting left.[19] Our result that in relatively rich societies it is actually the more equal income distributions which are less politically conducive for signalling is also consistent with the observation that rich US states are more likely to vote Democrat even though rich voters overall are more likely to vote Republican.[20]

We have made a few simplifying assumptions to facilitate our analysis. We analyze a simple signalling environment, which is one dimensional and characterized by a quasi-linear utility function. Our results can be extended to consider other utilities and more dimensions. We have provided general conditions for all forms of signalling, and a modelling of the supply side may provide more specific results. Most importantly, we have used a simple majority rule to assess the political viability of different policies. Clearly in some environments the median voter's preferences or those of the majority more generally are not sufficient to determine the political outcome. Organized lobbies which typically are more likely to represent organized high income voters or private providers of signals may bias the political outcome in their favour; these may imply that even if there are pressures for redistributions due to sorting, these are not necessary successful.

---

[19]See De La O and Rodden (2008).

[20]See http://redbluerichpoor.com/excerpt.php.

Appendix

Proof of Proposition 1: The necessary part follows from the case for one signal and $t = 0$. Note that with some tax level t, the condition becomes $\frac{(1-t)x+t\mu}{F(x)} \geq \mu$ and thus the condition for $t = 0$ is sufficient for any $t > 0$. This will also be the case for any FSP with $n > 2$ and thus we abstract away from taxation and set $t = 0$.

We now show sufficiency using an induction on the number of signals. We have already shown the sufficiency of the condition for $n = 2$. Suppose that the Proposition is true for any FSP with $n = k - 1$. Consider all FSP with $n = k$. Again we will focus on the utility of x^m, as the utility of all the types below are lower (and all the types above are higher) from the FSP.

Note that if $x^m < x_1$, then his utility is like in an FSP with $n = 2$ and the same x_1, and so Condition 1 applies. If $x_1 < x^m < x_2$, consider his utility from an FSP with $n = 3$ and the same x_1, x_2, which is the same again. Thus if $x_{i-3} < x^m < x_{i-2}$ for $i \leq k$, his utility from the FSP is the same as the utility from an FSP with $n = i$ and the same $x_0, x_1, ..., x_{i-2}$ which by the induction hypothesis proves the result. Now assume that $x_{k-2} < x^m < x_{k-1}$. His expected utility can be written as:

$$x_1 E(x_j | x_j \in [0, x_1]) + (x_2 - x_1) E(x_j | x_j \in [x_1, x_2]) + ...$$
$$+(x_{k-2} - x_{k-3}) E(x_j | x_j \in [x_{k-3}, x_{k-2}]) + (x^m - x_{k-2}) E(x_j | x_j \in [x_{k-2}, x_{k-1}])$$

which is strictly lower than the utility from an FSP with $n = k - 1$ and the same $x_0, x_1, ..., x_{k-2}$ in which case the last expectations are replaced by $E(x_j | x_j \in [x_{k-2}, 1])$ and the rest is the same.

Finally consider the case of $x^m > x_{k-1}$. We first divide and multiply his expected utility by x^m and then use Condition 1 repetitively:

$$x^m \left(\frac{x_1}{x^m} E(x_j | x_j \in [0, x_1]) + \frac{(x_2 - x_1)}{x^m} E(x_j | x_j \in [x_1, x_2]) + ... \right.$$
$$+ \frac{(x_{k-1} - x_{k-2})}{x^m} E(x_j | x_j \in [x_{k-2}, x_{k-1}]) + \frac{(x^m - x_{k-1})}{x^m} E(x_j | x_j \in [x_{k-1}, 1]) \leq$$
$$x^m \left(F(x_1) E(x_j | \in [0, x_1]) + (F(x_2) - F(x_1)) E(x_j | \in [x_1, x_2]) + ... \right.$$
$$+(F(x_{k-1}) - F(x_{k-2})) E(x_j | x_j \in [x_{k-2}, x_{k-1}]) + (1 - F(x_{k-1})) E(x_j | x_j \in [x_{k-1}, 1]) \leq x^m \mu$$

This completes the proof. Note that when $x^m = \mu$ we have Condition 1 whereas an analogous and stronger condition is when $\frac{x}{F(x)} \geq x^m$. ∎

Sufficiency of Condition 1 for other utility functions:

We now show that condition 1 is also sufficient for the utility function $(x-b)(y-b)$. Suppose

there is just one signal. Then sorting is better than FR for the median/mean if

$$(\mu - b)(\bar{E}_x - b) \leq \mu^2 \Leftrightarrow$$

$$F(x)(\bar{E}_x - \underline{E}_x) + \frac{b}{\mu}(b - \mu - \bar{E}_x) \leq 0$$

Note that at the cutoff x, we have that $(x - b)(\bar{E}_x - b) = x\underline{E}_x$, so that $x(\bar{E}_x - \underline{E}_x) = b(-b + x + \bar{E}_x)$. This implies that the above holds if

$$F(x)(\bar{E}_x - \underline{E}_x) \leq \frac{x(\bar{E}_x - \underline{E}_x)}{\mu}\frac{-b + \mu + \bar{E}_x}{-b + x + \bar{E}_x}$$

For which, as $\frac{-b + \mu + \bar{E}_x}{-b + x + \bar{E}_x} > 1$, a sufficient condition is Condition 1.∎

Proof of Proposition 2: Average utility from sorting for some FSP $\mathbf{x}$ can be written as

$$
\begin{aligned}
U(\mathbf{x}) &= F(x_0)E_0^2 + \sum_{i=1}^{n}(F(x_i) - F(x_{i-1}))E_i^2 + (1 - F(x_n))E_{n+1}^2 - \sum_{i=0}^{n}(1 - F(x_i))x_i(E_{i+1} - E_i) \\
&= \sum_{i=0}^{n}F(x_i)(E_i^2 - E_{i+1}^2) + E_{n+1}^2 - \sum_{i=0}^{n}(1 - F(x_i))x_i(E_{i+1} - E_i) \\
&= \sum_{i=0}^{n}(E_i - E_{i+1})[F(x_i)(E_i + E_{i+1} - x_i) + x_i] + E_{n+1}^2
\end{aligned}
$$

The average utility from full redistribution is:

$$
\begin{aligned}
U(FR) &= \mu(F(x_0)E_0 + \sum_{i=1}^{n}(F(x_i) - F(x_{i-1}))E_i + (1 - F(x_n))E_{n+1}) \\
&= \mu(\sum_{i=0}^{n}F(x_i)(E_i - E_{i+1}) + E_{n+1})
\end{aligned}
$$

Let $\Delta = U(\mathbf{x}) - U(FR)$.

$$\Delta = \sum_{i=0}^{n}(E_i - E_{i+1})[F(x_i)(E_i + E_{i+1} - x_i - \mu) + x_i] + E_{n+1}(E_{n+1} - \mu)$$

note that

$$E_{n+1} - \mu = E_{n+1} - \sum_{i=0}^{n}F(x_i)(E_i - E_{i+1}) - E_{n+1} = -\sum_{i=0}^{n}F(x_i)(E_i - E_{i+1})$$

Therefore,

$$\Delta = \sum_{i=0}^{n}(E_i - E_{i+1})[F(x_i)(E_i + E_{i+1} - x_i - \mu - E_{n+1}) + x_i]$$

17

Suppose that $n = 0$. Then

$$
\begin{aligned}
\Delta &= (E_0 - E_1)[F(x_0)(E_0 - x_0 - \mu) + x_0] < (>)0 \\
&\Leftrightarrow F(x_0)(E_0 - x_0 - \mu) + x_0 > (<)0 \\
&\Leftrightarrow F \text{ is } NBUE(NWUE)
\end{aligned}
$$

Suppose that $n = 1$. Then:

$$
\begin{aligned}
\Delta &= (E_0 - E_1)[F(x_0)(E_0 + E_1 - x_0 - \mu - E_2) + x_0] + \\
&\quad (E_1 - E_2)[F(x_1)(E_1 - x_1 - \mu) + x_1] \\
&= (E_0 - E_1)[F(x_0)(E_0 - x_0 - \mu) + x_0] + \\
&\quad (E_1 - E_2)[F(x_0)(E_0 - E_1) + F(x_1)(E_1 - x_1 - \mu) + x_1] \\
&= (E_0 - E_1)[F(x_0)(E_0 - x_0 - \mu) + x_0] + \\
&\quad (E_1 - E_2)[F(x_0)E_0 + E_1(F(x_1) - F(x_0)) - F(x_1)\mu + (1 - F(x_1))x_1] \\
&= (E_0 - E_1)[F(x_0)(E_0 - x_0 - \mu) + x_0] + \\
&\quad (E_1 - E_2)(1 - F(x_1))[\mu - \bar{E}_1 + x_1] \\
&< (>)0 \Leftrightarrow F \text{ is } NBUE(NWUE)
\end{aligned}
$$

Thus, for any n :

$$
\begin{aligned}
\Delta &= \sum_{i=0}^{n}(E_i - E_{i+1})[F(x_i)(E_i + E_{i+1} - x_i - \mu - E_{n+1}) + x_i] \\
&= \sum_{i=0}^{n}(E_i - E_{i+1})[F(x_i)(E_i + E_{i+1} + \sum_{i}^{n-2} E_{i+2} - \sum_{i=0}^{n-2} E_{i+2} - x_i - \mu - E_{n+1}) + x_i] \\
&= \sum_{i=0}^{n}(E_i - E_{i+1})[F(x_0)E_0 + \sum_{j=1}^{i}(F(x_j) - F(x_{j-1}))E_j - F(x_i)\mu + x_i(1 - F(x_i))] \\
&= \sum_{i=0}^{n}(E_i - E_{i+1})[\mu - (1 - F(x_i))\bar{E}_i - F(x_i)\mu + x_i(1 - F(x_i))] \\
&= \sum_{i=0}^{n}(E_i - E_{i+1})(1 - F(x_i))[\mu - \bar{E}_i - x_i]
\end{aligned}
$$

Where $\bar{E}_i$ is the expectations over $[x_i, \nu]$. Thus $\Delta < (>)0$ if F is NWUE (NBUE).∎

Proof of Lemma 1: (i) Note that when we write the utility from a general FSP as in Proposition 1, utility trivially increases in the expectations over types that one meets and

18

hence for any $z \in [x_i, x_{i+1}]$ utility is highest when $x_{i+1} = 1$. (ii) Suppose that the partition is $[0, x_i, x_{i+1}, ...]$ and that $z \in [x_i, x_{i+1}]$. The utility of z is

$$x_i \underline{E}_{x_i} + (z - x_i) E[x_j | x_j \in [x_i, x_{i+1}]] \tag{4}$$

Now suppose that the partition is $[0, x_{i-1}, x_i, x_{i+1}, ...]$. The utility of z is then:

$$x_{i-1} \underline{E}_{x_{i-1}} + (x_i - x_{i-1}) E[x_j | x_j \in [x_{i-1}, x_i]] + (z - x_i) E[x_j | x_j \in [x_i, x_{i+1}]] \tag{5}$$

The utility in (5) minus the utility in (4) is:

$$
\begin{aligned}
x_{i-1} \underline{E}_{x_{i-1}} + (x_i - x_{i-1}) E[x_j | x_j &\in [x_{i-1}, x_i] - x_i \underline{E}_{x_i} = \\
x_{i-1} \underline{E}_{x_{i-1}} + (x_i - x_{i-1}) E[x_j | x_j &\in [x_{i-1}, x_i] - x_i (\frac{F(x_{i-1})}{F(x_i)} \underline{E}_{x_{i-1}} \\
+(1 - \frac{F(x_{i-1})}{F(x_i)}) E[x_j | x_j &\in [x_{i-1}, x_i]) = \\
(x_{i-1} - x_i (\frac{F(x_{i-1})}{F(x_i)}) (\underline{E}_{x_{i-1}} - E[x_j | x_j &\in [x_{i-1}, x_i]) \geq 0 \Leftrightarrow \frac{x_{i-1}}{F(x_{i-1})} \leq \frac{x_i}{F(x_i)}
\end{aligned}
$$

Or in other words, $\frac{x}{F(x)}$ being an increasing function which is the case when F is concave. The argument above can be extended to any region of the partition which we can make finer and finer. Note that the cost of signals of intervals that converge to the point x converges to $\frac{x^2}{2}$. ∎

Proof of Proposition 3: Given Lemma 1 we need to show that the conditions above imply:

$$\frac{x^2}{2} + (x^m - x) \bar{E}_x \leq \mu^2$$

for any x, as any other FSP which includes this x as a border of an interval will provide a lower utility. Note further that we need to focus only on $x < x^m$. To see condition (i) above note that if we plug $\bar{E}_x \leq \frac{\mu^2}{x^m}$ then the above is lower than μ^2 whenever $x < x^m$. We therefore need $x^m > x > x^*$. To see condition (ii) note that for all $x < x^m$, $\bar{E}_x < 2\mu$ (by integration by parts and noting that $F(x) \leq \frac{1}{2}$), and hence the above decreases in x. At $x = 0$, the maximum utility from sorting at x is therefore $2x^m \mu \leq \mu^2$ iff $\frac{x^m}{\mu} \leq \frac{1}{2}$. To see condition (iii), note that the derivative of $\frac{x^2}{2} + (x^m - x) \bar{E}_x - \mu^2$ w.r.t. x is:

$$((x^m - x) \frac{f(x)}{1 - F(x)} - 1)(\bar{E}_x - x)$$

At x^*, $\frac{x^2}{2} + (x^m - x) \bar{E}_x - \mu^2 < 0$, and thus if $(x^m - x^*) \frac{f(x^*)}{1 - F(x^*)} - 1 < 0$ the derivative will be negative at x^* and will continue by DFR to be negative for all x implying the above. Extending

all the above arguments to allow for some $t > 0$ is straightforward and the conditions identified will be sufficient in this case as well.∎

Proof of Proposition 4: Consider expected utility from sorting for some type $z > x$:

$$(z - x)(1 - t)(\bar{E}_x(1 - t) + t\mu) + (x(1 - t) + t\mu)((1 - t)\underline{E}_x + t\mu) \tag{6}$$

For some type z. The derivative w.r.t. t is:

$$-(z - x)(\bar{E}_x(1 - t) + t\mu) + (z - x)(1 - t)(\mu - \bar{E}_x) +$$
$$(\mu - x)((1 - t)\underline{E}_x + t\mu) + (x(1 - t) + t\mu)(\mu - \underline{E}_x)$$
$$= (z - x)((1 - 2t)\mu - 2\bar{E}_x(1 - t)) +$$
$$(\mu - x)((1 - t)\underline{E}_x + t\mu) + (x(1 - t) + t\mu)(\mu - \underline{E}_x)$$

Note that the first (negative) element decreases with t whereas the other elements increase with t so for any x, preferences for redistribution if positive at some point must increase afterwards, but can be negative at first. Consider first $z = x^m = \mu$, which implies that the only motive is sorting and not redistribution. At $t = 0$ the above is:

$$\mu(\mu - \bar{E}_x) + (2x - \mu)(\bar{E}_x - \underline{E}_x)$$

which implies that for low enough x, preferences are first decreasing. Around $x = \mu$ preferences are always positive. Around t close to 1:

$$(\mu - \underline{E}_x)\mu > 0$$

so preferences for t must increase at the end. In that case, the effect on the type that one meets is too low as well as the information rent.

If we add redistribution motive then note that the worst point is $x = 0$, $t = 0$. Then also for $x^m < \mu$ we have preferences against redistribution as it comes to

$$\mu\underline{E}_x + \mu x^m - 2x^m\bar{E}_x < 0$$

as $\underline{E}_x < x^m$. Of course if x^m is not in the club he likes redistribution and around t close to 1 we have as above.∎

References

[1] Bagwell, L. S. and B.D. Bernheim (1996), "Veblen Effects in a Theory of Conspicuous Consumption," American Economic Review, vol. 86(3), pp 349-73.

[2] Barlow, R. E. and F. Proschan (1965), Mathematical Theory of Reliability (New York: Wiley).

[3] Barlow, R. E. and F. Proschan (1966), "Inequalities for Linear Combinations of Order Statistics from Restricted Families", Annals of Mathematical Statistics, vol. 37, pp. 1593–1601.

[4] Benabou, R. (1996), "Heterogeneity, Stratification and Growth: Macroeconomic Implications of Community Structure and School Finance", American Economic Review, 86, pp. 584-609.

[5] Cramer, J. S. (1978), "Function for Size Distribution of Incomes: Comment", Econometrica Vol. 46, pp. 459-460.

[6] Damiano, E. and Li, H. (2007): "Price discrimination and efficient matching", Economic Theory, vol. 30, 243-263.

[7] De La O, A.L and J.A. Rodden (2008), "Does Religion Distract the Poor? Income and Issue Voting Around the World", Comparative Political Studies, vol. 41 pp. 437-476.

[8] Epple, D. and R. Romano (1998), "Competition between Private and Public Schools, Vouchers, and Peer-Group Effects", American Economic Review, vol. 88, pp. 33-62.

[9] Fernandez, R. and R. Richardson (2003), "Equity and Resources: An Analysis of Education Finance Systems" , Journal of Political Economy, pp.

[10] Fernandez, R. and R. Richardson (2001), "Sorting and Long-Run Inequality" Quarterly Journal of Economics,

[11] Fernandez. R. and J. Gali, J. (1997), "To Each According to...? Markets, Tournaments and the Matching Problem with Borrowing Constraints", Review of Economic Studies, 66, 799–824.

[12] Glazer, A. and K.A. Konrad (1996), "A Signaling Explanation for Charity," American Economic Review, vol. 86(4), pp. 1019-28

[13] Hall, W. J. and Wellner, J. A. (1984), Mean residual life, In Proceedings of the international symposium in Statistics and Related Topics, pp. 169-184. Amsterdam: North-Holland

[14] Heffetz, O. (2011), "A Test of Conspicuous Consumption: Visibility and Income Elasticities", Review of Economics and Statistics, Vol. 93, pp. 1101-1117.

[15] Hoppe, H., B. Moldovanu, and A. Sela (2009), "The Theory of Assortative Matching Based on Costly Signals", Review of Economic Studies, vol. 76, p. 253-281

[16] Hoppe, H., B. Moldovanu, and E. Ozdenoren (forthcoming), "Coarse Matching with Incomplete Information", Economic Theory.

[17] Leibenstein, H. (1950), "Bandwagon, snob and Veblen effects in the theory of consumer demand", Quarterly Journal of Economics 64, 183-207.

[18] McAfee, R. P. (2002), "Coarse matching", Econometrica 70, pp. 2025-2034.

[19] Meltzer, A. and S. Richards (1981), "A rational theory of the size of government", Journal of Political Economy vol. 89, pp. 914-927.

[20] Moav, O. and Z. Neeman (2010), "Status and poversty", Journal of the European Economic Association, vol. 8, pp. 413-420.

[21] Pesendorfer, W. (1995), "Design Innovations and Fashion Cycles", American Economic Review, vol. 85, pp. 771–792.

[22] Rayo, L. (2005), "Monopolistic Signal Provision", B.E. Journal in Theoretical Economics.

[23] Rege, M. (2003), "Why Do People Care About Social Status", Journal of Economic Behavior & Organization, 66, pp. 233–242.

[24] Roemer, J. (1998), "Why the poor do not expropriate the rich", Journal of Public Economics, 70, pp. 399-424.

[25] Salem, A. B. Z. and T. D. Mount (1974), "A Convenient Descriptive Model of Income Distribution: The Gamma Density", Econometrica, Vol. 42, pp. 1115-1127.

[26] Singh S. K. and G. S. Maddala (1976), "A Function for Size Distribution of Incomes", Econometrica, Vol. 44, pp. 963-970.